Journal of Biomolecular Techniques
11:74–86 © 2000 ABRF

Systematic Analysis of Peptide Recoveries from In-Gel Digestions for Protein Identifications in Proteome Studies

Kaye D. Speicher, Olivera Kolbas, Sandra Harper, and David W. Speicher

The Wistar Institute, Philadelphia, PA

Metabolically radiolabeled recombinant proteins were used to systematically evaluate peptide recoveries from in-gel trypsin digestion. At least 80% of the labeled tryptic peptides could be readily extracted from gel bands containing 1 to 10 pmol, and at least 70% could be extracted at 200- to 500-fmol levels using a recombinant 52-kd protein. Alkylation before electrophoresis or before trypsin digestion had minimal effects on peptide recovery; although alkylation, especially before gel analysis, may reduce heterogeneity of resulting peptides containing cysteines. Comparison of different gel thicknesses using unminced gel bands suggested that 1.0-mm gels were optimal. Surprisingly, peptide recoveries from 0.5-mm gels were low and variable, primarily because of increased diffusion of protein out of thin gels during fixing and staining. Although 70% to 85% of tryptic peptides could typically be extracted from gels over a range of conditions and protein concentrations, further processing of peptide extracts resulted in substantial additional losses. Even minimal handling resulted in loss of about 10% to 15% of extracted peptides by adsorption to plastic surfaces. Adsorptive losses were particularly high, sometimes exceeding 50%, and variable if extracts were partially dried in a Speedvac to concentrate the sample or to remove acetonitrile. High acetonitrile extraction and/or Speedvac concentration appear to be detrimental, and their elimination simplifies sample handling and automation. SYPRO Ruby Red, a sensitive noncovalent fluorescent stain appears to be an attractive alternative to Coomassie blue for in-gel trypsin digestion. These results suggest an optimized in-gel trypsin digestion strategy in which proteins in 1.0-mm-thick gels are stained with Coomassie blue or Ruby Red, digested overnight with modified trypsin, and extracted one or two times with small volumes of aqueous buffer. It is especially critical that subsequent surface exposure be minimized, and concentration by vacuum drying should be avoided. (J Biomol Tech 2000;11:74–86)

KEY WORDS: in-gel digestion, protein identification, proteome analysis, tryptic peptides, peptide adsorption.

ADDRESS CORRESPONDENCE AND REPRINT REQUESTS TO: David W. Speicher, The Wistar Institute, Room 151, 3601 Spruce Street, Philadelphia, PA 19104 (email: speicher@wistar.upenn.edu).

Protein characterization tools have steadily evolved and improved in sensitivity to the point where the most practical method for isolating the final protein sample for structural analyses or identification is usually one-dimensional (1D) or two-dimensional (2D) polyacrylamide gel electrophoresis (PAGE). In addition, explosive interest in analyzing the entire protein component (ie, proteome) of cells and tissues has further expanded interest in reliable high-sensitivity identifications of proteins isolated by 2D gels.[1–4] Early reliable methods for characterizing proteins from gels involved electroblotting the proteins onto either nitrocellulose for subsequent tryptic digestion[5] or polyvinylidene difluoride (PVDF) for direct N-terminal sequencing.[6] Subsequently, methods were developed for in situ digestion of proteins in Coomassie blue–stained gels or bound to PVDF membranes,[7–10] and proteolysis of more than 10 pmol of proteins in gels or on membranes for internal automated sequencing became fairly routine for most laboratories.[11,12]

The increased emphasis within the past several years on identification of proteins at sub-picomole levels using mass spectrometry–based methods has shifted in situ proteolysis preferences to in-gel methods rather than membrane-based methods.[13,14] Current mass spectrometers are capable of identifying proteins by analysis of proteolytic fragments when low femtomole to attomole amounts are introduced into the instrument. However, the practical sensitivity for identification of proteins in gels is much less than the mass

spectrometry sensitivity because only a small portion of a larger-scale in-gel protease digestion is usually analyzed. During the past several years, numerous variations of the original in-gel digestion methods[8,10] have been developed by many laboratories, including miniaturized and automated methods.[15,16]

Despite the key role of in-gel protease digestion in most proteome research strategies, the influence of modifying different steps of the procedure on peptide yields is poorly understood. Systematic evaluation and optimization of in-gel digestion using mass spectrometry (MS) as the detection method is complicated by the fact that MS methods are not intrinsically quantitative and are sensitive to signal suppression by many contaminants, particularly near the instrument sensitivity limits. Even repetitive analyses performed on a single-peptide mixture can yield substantial variations in peptide mass signal strengths and the number of peptides detected, particularly when matrix-assisted laser desorption and ionization (MALDI) MS is used. Hence, modifications of in-gel digestion and sample processing steps that improve MS signals may represent either improved peptide yields or simply removal of suppressive components. More importantly, such end-point analyses cannot directly indicate peptide recoveries at individual steps of the procedure. Although enhancement of MS signals is the ultimate goal regardless of how it is accomplished, separate evaluation of peptide yields and contaminant removal should enable rational selection of the best conditions throughout each step of the overall method.

In this study, we used metabolically radiolabeled recombinant proteins to systematically evaluate peptide recoveries from in-gel trypsin digestion. Although our goal is to optimize in-gel digestion at femtomole levels, we started our evaluation at the 10-pmol level for several reasons. First, previous studies showed that, when many laboratories were surveyed, the yields of tryptic peptides as determined by initial Edman sequence yields were typically only 10% to 30% when 10 to 20 pmol of a protein were present in a gel band,[11,12] suggesting that further optimization at picomole levels may be feasible. Second, the 10-pmol level served as baseline for comparison with smaller protein amounts. Losses at each step of the method were quantitatively evaluated, and the effects of several variations, including alkylation, gel thickness, gel volume, gel type, postextraction processing, and use of a high-sensitivity fluorescent stain, were evaluated.

MATERIALS AND METHODS

A recombinant 78-kd fusion protein containing a 26-kd N-terminal GST domain and a 52-kd domain of human red cell spectrin was expressed in *Escherichia coli* as previously described[17] with modifications. Briefly, an overnight culture of GSTα18–21 was diluted 1:10 into 60-mL cysteine- and methionine-free minimal media and grown at 30°C. At 1 hour before induction with 1 mM (final concentration) of isopropyl-β-thiogalactopyranoside (optical density 0.5 at 550 nm), 10 mCi Pro-mix L-[^{35}S-cysteine/methionine] in vivo cell labeling mix (Amersham Pharmacia Biotech) was added to the culture. After induction, the culture was allowed to grow an additional 3 hours, and bacteria were collected by centrifugation. The fusion protein was purified as follows. The cell pellet was resuspended in 3 mL of lysis buffer, lysed by sonication at 0°C, and centrifuged. The supernatant was incubated with reduced glutathione Sepharose 4B for 1 hour at 4°C; the resin was washed with PBS, the fusion protein was eluted with G-buffer (50 mM Tris, 10 mM reduced glutathione, 1 mM 2-mercaptoethanol, 1 μg/mL pepstatin, 5 mM EDTA, pH 8.0); a portion of the sample was digested with thrombin; thrombin was inactivated by addition of 3 μM phenylmethylsulfonyl fluoride (PMSF); and the digested sample was concentrated with an Amicon 10-kd MWCO centrifugal concentrator and purified on two analytical high-performance liquid chromatography (HPLC) gel filtration columns (TSK G3000 SW_{XL} + G2000SW_{XL}, 7.8 × 300 mm each) in series at 0.8 mL/min in PBS (10 mM phosphate, 130 mM NaCl, 1 mM EDTA, 0.15 mM PMSF, 1 mM 2-mercaptoethanol, pH 7.4). The ^{35}S-α18–21 protein and ^{35}S-GST protein were separately pooled, and the purified radiolabeled proteins were stored as frozen aliquots before use.

Four gels were typically run at one time using a single radiolabeled protein aliquot, which was initially mixed with an equal volume of 2× SDS-PAGE sample buffer, diluted to 0.5 pmol/μL with 1× sample buffer, and heated to 37°C for 15 minutes. In most experiments, further dilutions were then prepared as needed using 1× sample buffer such that volumes loaded onto SDS-PAGE gels were between 4 and 20 μL to minimize errors in amounts applied to the gel. Typically, 8% Tricine[18] mini-gels (1.0 × 50 × 85-mm separation gel) were used. Gels were prepared 24 hours before use and stored at room temperature to ensure complete polymerization and to allow time for free radicals to decay. Thioglycolate (0.1 mM) was added to the top electrode buffer and electrophoresis was terminated when the bromophenol blue tracking dye was approximately 5 mm from the bottom of the gel. In some experiments, samples were alkylated before SDS-PAGE as follows. The 20× reducing buffer (500 mM Tris, 10% SDS, 20 mM tris[2-carboxyethyl]phosphine hydrochloride [TCEP-HCl], pH 8.0) was diluted 20-fold into the sample and incubated for

10 minutes at 37°C. Then 20× alkylation reagent (100 mM iodoacetamide, pH 8) was added and incubated for an additional 15 minutes at 37°C. The 5× gel sample buffer was added, and the sample was incubated for another 15 minutes at 37°C. Typically, gels were stained with 0.05% Coomassie blue R250 in 10% acetic acid and 50% methanol for 1 hour, destained for 2 to 4 hours in a solution containing 10% acetic acid and 15% methanol, and rinsed with water for 1 hour. Images were acquired using a digital camera (Alpha Innotech Corp., Alexandria, VA) and samples were stored at 4°C. Some gels were stained with 50 mL of SYPRO Ruby Red (Molecular Probes, Eugene, OR) for 3 to 16 hours and destained in water for 16 to 24 hours; images were acquired with a laser densitometer, and gels were stored at 4°C in the dark.

The consistency of sample loading of replicate lanes and different amounts was verified by densitometry of stained gels. The amount of protein in the gel bands was initially determined by amino acid analysis of several gel bands, and resulting values were correlated with densitometry. Thereafter, concentrations of new samples were verified by densitometry alone using identical conditions. These values were typically about 5% to 15% lower than the concentrations calculated from absorbances at 280 nm, which was consistent with the apparent purity of the proteins (ie, 85% to 95% pure). Protein recoveries in gel bands and the amount of radiolabel covalently associated with these bands were determined by scintillation counting. Selected lanes were excised and placed in scintillation vials; 300 μL of 1 M NaOH was added; and samples were incubated for 3 hours at 60°C, cooled, and neutralized with 27 μL of concentrated HCl. Then 4 mL of scintillation fluid was added, and radioactivity was measured in a scintillation counter.

The basic trypsin digestion method used in this study is a modification of the method originally described by Hellman et al.[10] All plastic microfuge tubes were precleaned by rinsing several times with 0.1% trifluoroacetic acid (TFA) and 50% acetonitrile before use as digestion or sample tubes. Protein bands were excised from stained SDS gels using a clean, sharp scalpel at the margin of detectable stain and transferred to 500-μL microfuge tubes. The gel volume was measured; bands were destained two times using 200 μL of 200 mM ammonium bicarbonate and 50% acetonitrile for 45 minutes at 37°C; and gel pieces were dried using a Speedvac. In some experiments, samples were alkylated before trypsin digestion as follows: 100 μL of 2 mM TCEP in 25 mM ammonium bicarbonate (pH 8.0) was added to the dried gel and incubated for 15 minutes at 37°C with agitation; the supernatant was then removed; 100 μL of 20 mM iodoacetamide in 25 mM ammonium bicarbonate (pH 8.0) was added and incubated for 30 minutes at 37°C in the dark; the supernatant was then discarded; the gel band was washed three times with 200 μL of 25 mM ammonium bicarbonate for 15 minutes, each with agitation; and gel bands were dried in a Speedvac. Dried gel bands with or without prior alkylation were rehydrated with 20 μL, or 1.5 times the gel volume if greater than 20 μL, of 0.02 μg/μL Promega-modified trypsin in 40 mM ammonium bicarbonate with 10% acetonitrile (pH 8.1) for 1 hour at room temperature. An additional 50 μL of 40 mM ammonium bicarbonate and 10% acetonitrile was then added, and digestion was continued for 16 to 18 hours at 37°C with agitation. After incubation, the supernatant was removed (extract 1). Another 50 μL of 0.1% TFA was added and incubated for 45 minutes at 37°C with agitation, and the supernatant was removed (extract 2). Another extraction was performed using the same conditions (extract 3).

In most experiments, recoveries and losses at each step of the digestion process were measured. Liquid fractions such as destain solution and extracts were transferred directly to scintillation vials. Solid surfaces such as pipet tips and microfuge tubes used for the digestion were extracted two times with 200 μL of 0.1% SDS, and the extracts were combined in a scintillation vial. Digested gel pieces were hydrolyzed and counted as described previously for undigested gel bands.

RESULTS

Selection of Protein Standards and Consistency of Peptide Recoveries

A 78-kd recombinant fusion protein was metabolically radiolabeled to high specific activity on methionines and cysteines for this study. The use of a metabolically labeled protein ensured that the counts per minute (cpm) measured would reflect all molecules in the sample. This approach was more reliable than the alternative of mixing cold protein with a tracer amount of heavily iodinated protein, which could often behave quite differently because of extensive chemical modification. The location of a thrombin cleavage site between the two domains of the fusion protein allowed the facile isolation of two unrelated proteins with different characteristics from a single radiolabeling experiment. The 26.2-kd GST moiety has a predicted pI of 6.1 and 13 radiolabeled residues (9 Met, 4 Cys) in 7 tryptic peptides ranging in mass from 697 to 4059 daltons. This domain was used in some experiments as a representative small protein. The 52-kd spectrin domain was selected for this study from more than 30 GST fusion proteins available in the laboratory for several reasons. Its 51.9-kd mass and pI of 4.7

TABLE 1

Predicted Tryptic Peptides Larger than 700 Daltons from the 52-kd Protein[a]

Residues	Monoisotopic Mass	HPLC	Sequence
389–394	718.37	11.90	EIQA**M**K
219–224	725.44	14.78	QLPLQK
326–331	773.39	10.43	EQELQK
87–93	815.44	11.47	IEALNEK
169–176	927.53	17.17	LPEITDLK
73–81	1082.54	10.20	VLHEESQNK
317–325	1110.57	17.56	HLSDIIEER
179–188	1124.59	13.35	LISAQHNQSK
362–371	1125.61	22.25	AYFLDGSLLK
236–245	1134.51	16.49	ASALNNW**C**EK
285–294	1201.64	22.86	**C**LLELDQQIK
121–131	1215.61	16.91	ADVVEAWIADK
268–278	1272.61	17.90	DHEDFLASLAR
225–235	1304.64	18.96	AEDLFVEFAHK
372–384	1418.69	13.67	ETGTLESQLEANK
201–212	1437.74	20.40	WEQLLEASAVHR
59–72	1529.75	16.67	VQNV**C**AQGEDILNK
29–44	1622.76	21.42	GD**C**GDTLAATQSLL**M**K
45–58	1666.77	15.48	HEALENDFAVHETR
137–153	1718.88	20.07	TNGNGADLGDFLTLLAK
154–168	1764.83	19.20	QDTLDASLQSFQQER
436–450	1862.92	21.75	**M**QHNLEQQIYPGIHR
400–416	1868.01	24.70	IVDLGDNLEDALILDIK
106–120	1944.89	25.05	LQLEDDYAFQEFNWK
246–263	2097.98	21.44	**M**EENLSEPVH**C**VSLNEIR
295–313	2116.12	25.48	ALGVPSSPYTWLTVEVLER
417–435	2252.16	29.10	YSTIGLAQQWDQLYQLGLR
1–22	2657.20	25.53	GSLEYLQF**M**QNAEEEEAWINEK
339–361	2863.29	29.92	NFE**MC**QEFEQNASTFLQWILETR

[a]The 17 tryptic peptides that are 2 to 6 residues long and less than 700 daltons are not shown because none of these peptides contains a radiolabeled residue. Peptide masses and HPLC elution times are from the computer program GPMAW. The 6 cysteines and 6 methionines that are metabolically radiolabeled are indicated with bold underlined type.

places it within the most densely populated region of most 2D gel patterns for whole-cell extracts, making it a good mid-range representative protein. More importantly, the predicted radiolabeled peptides provided an excellent test group for evaluating in-gel digestions (Table 1). The N-terminal peptide and the peptide within one residue of the C-terminus were radiolabeled, and the remaining seven labeled peptides were spread throughout the remainder of the protein. This ensured that good peptide recovery would require efficient trypsin cleavage of all regions of the protein. All labeled peptides were larger than 700 daltons, and 9 of 12 labeled residues (6 Cys, 6 Met) were in peptides that were larger than 1500 daltons. This size distribution covers the optimal range for both peptide mass measurement and utility for protein identification, and the absence of small labeled peptides avoided any bias for nonrepresentatively easy extraction that small labeled peptides would have provided. The predicted HPLC elution positions were distributed throughout most of the elution range typically represented by tryptic peptide maps. The tryptic peptides produced by the 52-kd protein were therefore highly representative of an average protein, and this protein was used for most experiments.

Typically, sets of four 1.0-mm-thick gels were run in parallel, starting with a single protein stock solution that was used for serial dilutions in the presence of SDS to avoid adsorptive losses. In most experiments, radioactivity at each step of the trypsin digestion method was determined by transferring all solutions directly to scintillation vials, extracting all sample

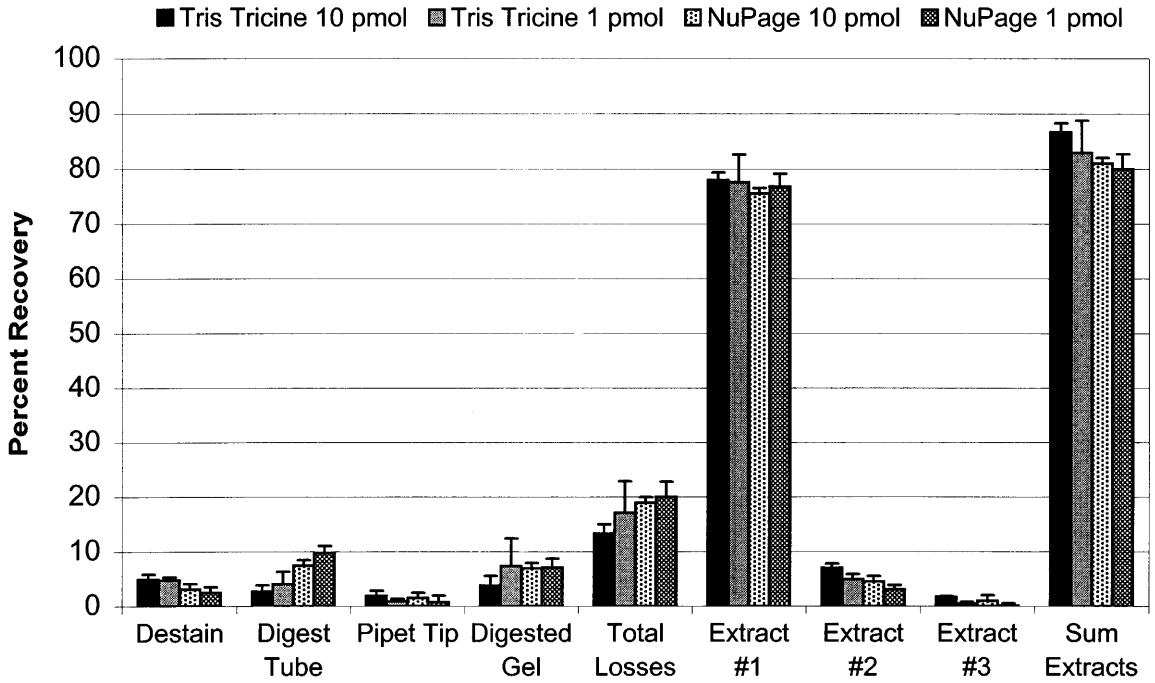

FIGURE 1

Peptide recoveries using the standard trypsin digestion and extractions. Coomassie blue–stained bands of the 52-kd recombinant (10- and 1-pmol loads) were excised from 8% Tricine gels and 4% to 12% NuPAGE gels with MES-SDS buffer (Invitrogen/NOVEX, Carlsbad, CA), destained, and digested with trypsin. Recoveries were determined as described in the Methods section. All extractions were transferred directly to scintillation vials to minimize adsorptive losses. The values for pipet tip represent counts recovered from the single pipet tip used for transferring the three sequential extracts. The average of at least two samples for each condition and standard deviations are shown.

contact surfaces with SDS as described previously, and hydrolysis of the digested gel. When all counts were summed for an entire digested sample, the recovered radioactivity was usually within ±5% of the counts recovered from duplicate undigested bands containing more than 1 pmol and within ±10% of undigested gel bands containing less than 1 pmol (data not shown). Because most of these minor variations apparently reflected pipetting precision while loading samples to gels, percentage recoveries were reported based on the summed total of all counts recovered for each sample when the total recovered counts were within the above error limits. Exceptions to this approach, which include experiments in which all possible sources of losses were not measured, are noted subsequently.

Peptide Recoveries Using Standard Trypsin Digestion and Extractions

The results of a representative experiment in which losses were evaluated at each step using our normal trypsin digestion and extraction procedure as described in the Methods section are shown in Figure 1. There were no substantial differences between 1- and 10-pmol loads or between bands from Tricine and NuPAGE gels. In all cases, most of the total radioactivity was recovered in the first extract (>75%), with minor amounts in the second extraction and negligible amounts in the third extraction. Although the total losses added up to 14% to 20%, losses at any single step were fairly minor, suggesting that further improvements in peptide extraction were unlikely to be feasible. One important point is that this experiment was designed to evaluate how effectively peptides could be extracted. Because these extracts were placed directly into scintillation vials for accurate quantification, subsequent adsorptive losses due to normal downstream handling are not represented. Hence, the actual amount of peptides that would be recovered in an actual protein identification experiment should be less than the summed extracts shown here because of subsequent adsorptive losses. Downstream adsorptive losses of extracted peptides were evaluated in subsequent experiments. In addition to potential downstream adsorptive losses, about 25% to 50% of the total

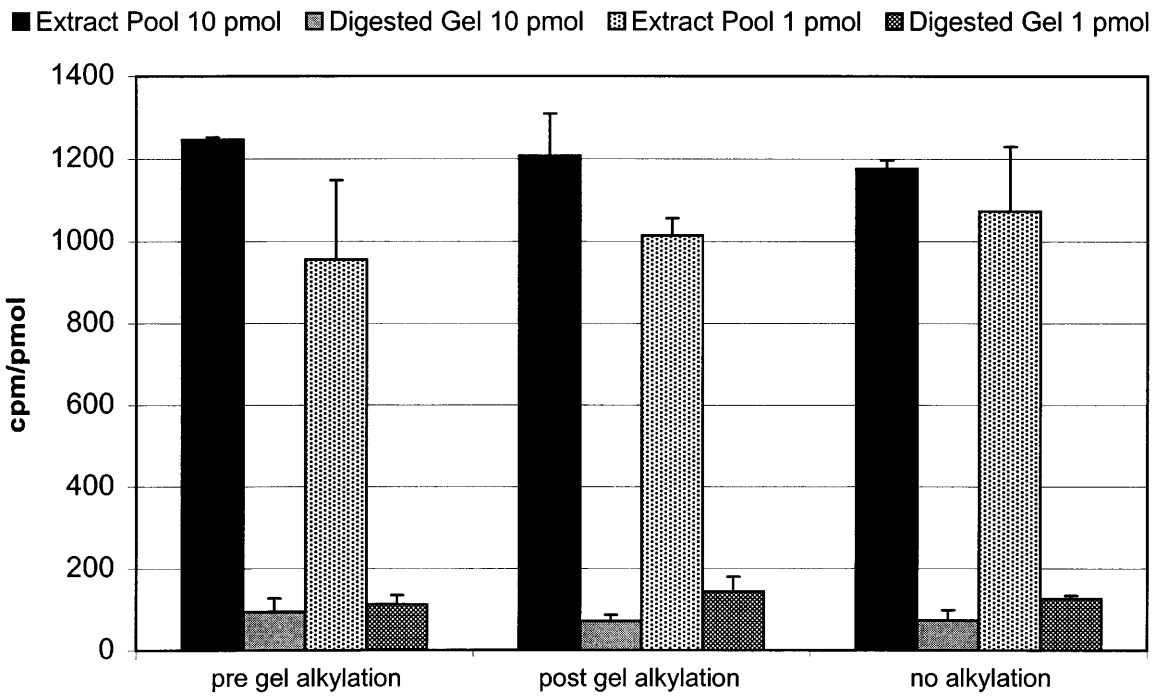

FIGURE 2

Effect of alkylation on peptide recoveries. The 52-kd protein was excised from 8% Tricine gels stained with Coomassie blue (10- and 1-pmol loads). The effects of alkylation before gel electrophoresis and alkylation after gel electrophoresis were compared with parallel samples digested without cysteine modification. Because not all fractions of the digestion protocol were counted, average cpm/pmol values are reported to aid direct comparison of the two protein amounts. The average of at least two samples for each condition and standard deviations are shown.

losses in the experiment shown in Figure 1 are adsorptive losses (ie, the microfuge tube used for the digestion and the pipet tip used to transfer extracts to scintillation vials). The small but consistent losses in the destain solution are apparently caused by proteins near the surface of the gel that initially diffuse out of the gel before the pores shrink in the 50% acetonitrile, because most of the radioactivity eluted at this step was recovered in the first of the two washes (data not shown).

Effect of Alkylation on Peptide Recoveries

Most protease digestion protocols recommend alkylation of the protein before proteolysis to help maintain a denatured protease-sensitive state and to avoid formation of mixed disulfide bonds, which introduces undesirable mass and peptide heterogeneity. A recent study showed that alkylation of protein samples before SDS-PAGE can minimize partial modification of cysteine containing peptides during electrophoresis.[19] To evaluate the effects of alkylation on peptide recoveries (Fig. 2), a stock solution of the 52-kd recombinant protein was divided into two parts, and one aliquot was modified with iodoacetamide before replicate loading of both aliquots onto a set of gels. In this experiment, similar amounts of radiolabel were recovered for all conditions, although there was some variation between experiments (ie, sometimes alkylated samples showed slightly higher yields than unalkylated samples).

The effect of alkylation on peptide heterogeneity was evaluated by separating tryptic peptides from 10-pmol digestions using reverse-phase HPLC (Fig. 3). In this experiment, a slightly higher yield of peptides was observed for the alkylated sample based on radioactivity, and this increased yield was confirmed by higher absorbances for most peaks in the alkylated sample. Surprisingly, there was no noticeable difference in heterogeneity of peptides that contained cysteine because the distribution of radiolabel in the two chromatographs was very similar. In addition to the major peaks containing most of the radiolabel, numerous very minor peaks contained low levels of radiolabel, which is indicative of peptide heterogeneity that is likely to be a combination of incomplete cleavages and other microheterogeneity of the peptides, including acrylamide modification and oxidation. This substantial peptide heterogeneity is probably a major

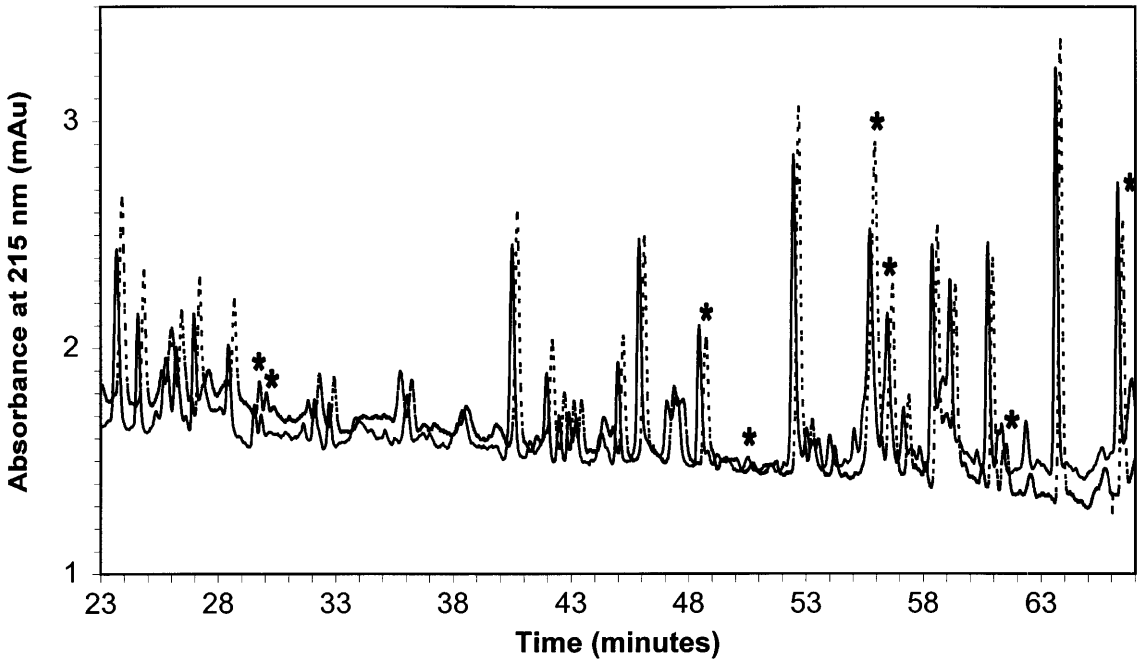

FIGURE 3

Comparison of HPLC tryptic peptide maps from alkylated and nonalkylated 52-kd recombinant proteins. The 52-kd protein (10 pmol) was alkylated after gel electrophoresis (*dashed line*) or not modified before in-gel trypsin digestion (*solid line*). The entire pooled extract was then injected onto a 2.1 × 150 mm Zorbax 300 SB-C18 column (Hewlett Packard, Palo Alto, CA) and eluted at 200 μL/min using a 0.1% TFA and acetonitrile gradient. Peaks were collected directly into scintillation vials using an Isco peak separator and FOXY fraction collector. The portion of the chromatograph containing peptide peaks is shown, and peaks containing more than 150 cpm are labeled with asterisks.

contributor, together with downstream adsorptive losses, to the relatively low and variable Edman sequencing yields typically observed at the 10-pmol protein level[11,12] despite overall highly efficient extraction of labeled peptides, as indicated previously.

Effects of Gel Thickness and Volume on Peptide Recovery

Different thicknesses of gels and different widths of sample wells were used to produce protein samples with equal quantities distributed in the gel matrix at different densities to evaluate the effects of gel volume on peptide recoveries. As discussed earlier, gel loading consistency was usually verified within each series of gels using densitometry and by counting several undigested gel bands. However, densitometry was not necessarily expected to be consistent when different gel thicknesses were used because of potential variations in staining efficiency. Representative results from analysis of undigested bands from several experiments are shown in Figure 4. Radiolabel recoveries were consistent at all protein loading amounts when 1.0-mm gels were used. The increased data scatter at the 1.0- and 0.5-pmol protein loads probably resulted from pipetting variations, because in these early experiments all lanes were loaded from a single stock protein solution at 0.5 pmol/μL. As a result of this observed variability, serial dilutions of the stock solution were performed as needed to ensure that no less than 4 μL was loaded per lane in subsequent experiments. Surprisingly, recoveries from 0.5-mm gels were lower than from 1.0-mm gels despite use of a single stock solution for loading these gels.

Peptide recoveries using different gel thicknesses and volumes are shown in Figure 5. These data are illustrated using cpm/pmol units, because variable protein recoveries in gel bands were observed for different gel thicknesses. The best peptide recoveries were obtained with 1.0-mm gels. Most of the reduced yield in the thin gels was apparently caused by increased losses during staining. Not surprisingly, more counts were extracted in the second and third extractions when 1.5-mm-thick, larger-volume bands were used.

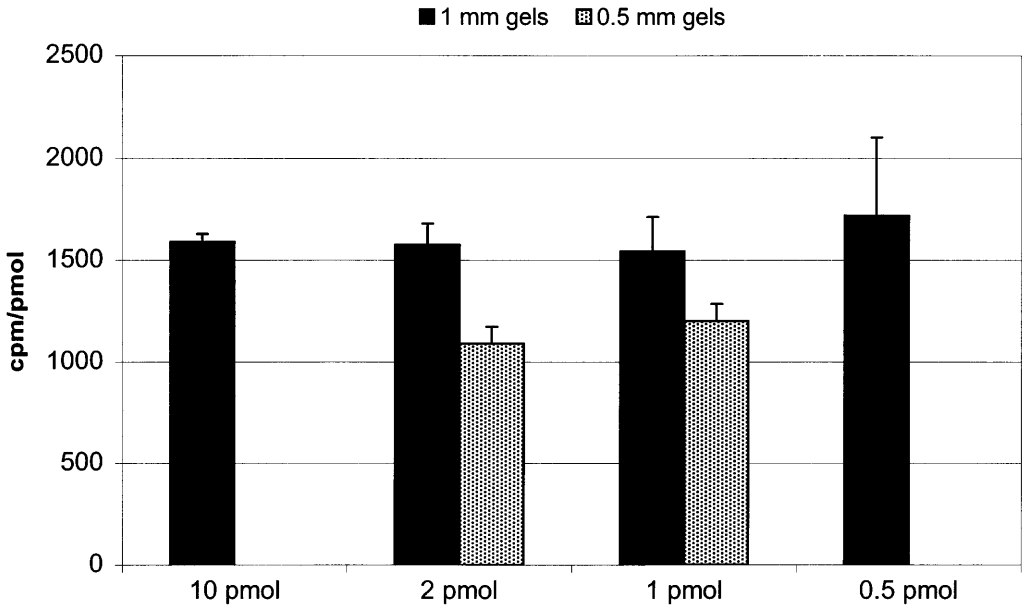

FIGURE 4

Reproducibility of gel bands. Replicate undigested gel bands (n = 2 to 6) from a typical set of gels were hydrolyzed and counted.

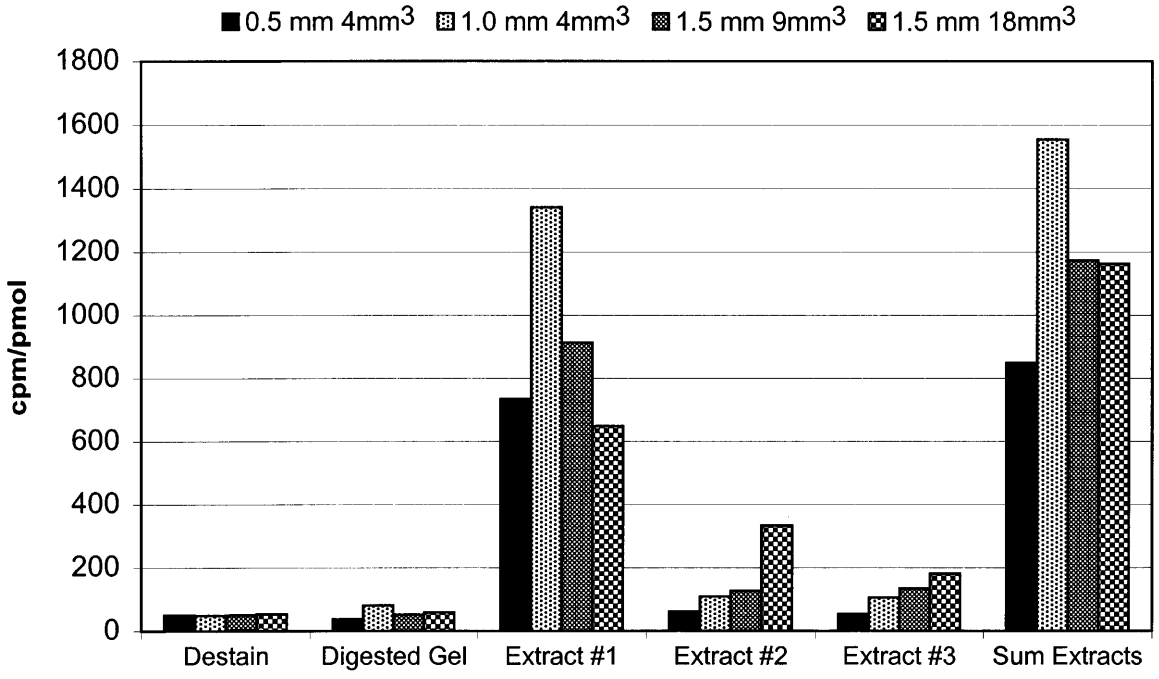

FIGURE 5

Effects of gel thickness and volume on peptide recoveries. Trypsin digestions of replicate bands in different thicknesses of gels were performed as described in Figure 1 and the Methods section. Values are reported in cpm/pmol units rather than percentages because not all possible sources of losses were measured and there were substantial variations in the amount of protein in bands from the different thicknesses of gels.

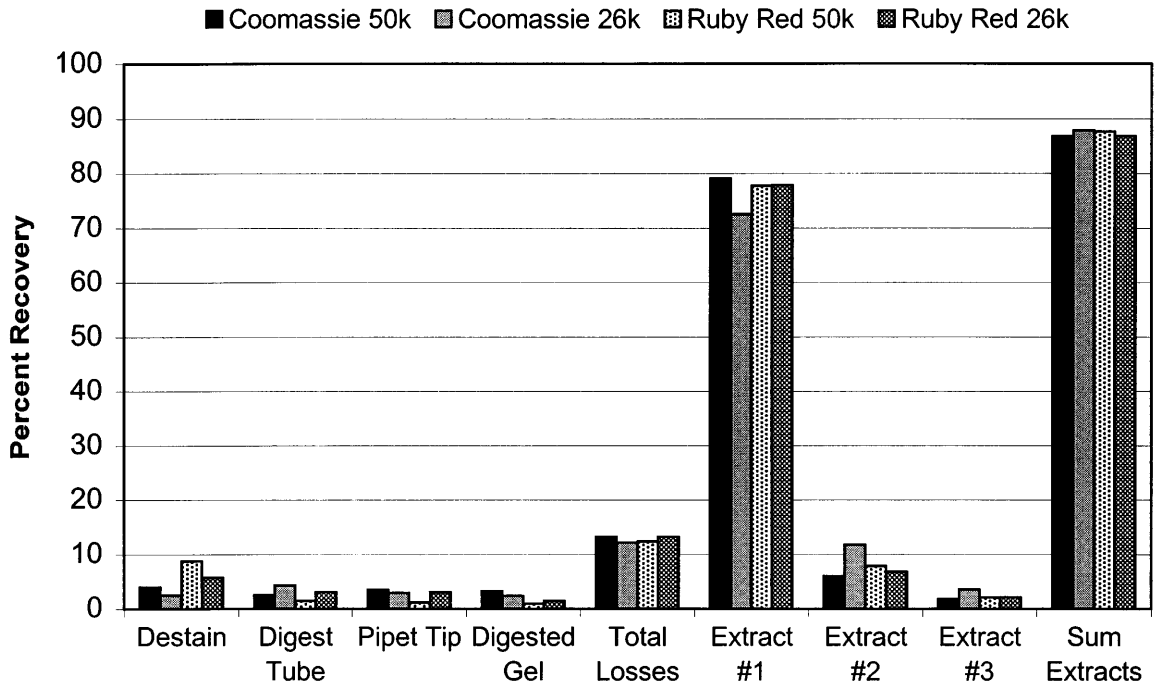

FIGURE 6

Comparison of different proteins and stains. Trypsin digestions were performed on 52-kd spectrin domain and 26-kd GST protein bands excised from gels stained with Coomassie blue or SYPRO Ruby Red. Recoveries were calculated as described in Figure 1 and the Methods section.

Comparison of Different Proteins and Stains

Parallel trypsin digestions were performed on the 52-kd spectrin domain and 26-kd GST protein to evaluate whether peptide recoveries observed for the 52-kd protein were representative of other proteins (Fig. 6). In addition, the effects of a promising high-sensitivity noncovalent fluorescent stain, SYPRO Ruby Red,[20] were evaluated using both proteins. Total peptide recoveries were high and essentially identical for all conditions, and only subtle differences occurred at several individual steps. For example, the protein losses during the destaining steps were more than twice as high for both proteins when Ruby Red was used. Although these losses remained relatively minor, higher destain losses were consistently seen in subsequent experiments using Ruby Red (data not shown). This consistent trend apparently reflects less effective fixation of the proteins in the gel compared with the Coomassie blue staining procedure, despite our use of a 1-hour fixation in a solution containing 10% acetic acid and 50% methanol. These increased losses suggest that the manufacturers recommendations of either fixation as optional or use of mild fixation conditions when fixation is desired may not be optimal for in-gel digestion when maximal overall recoveries are desirable. Alternatively, perhaps a more rapid gel dehydration step could be used for Ruby Red, because destaining per se is presumably not needed. In other experiments, use of Tris-glycine gels, as well as analysis of unstained gel bands located by position relative to bracketing prestained standards, gave similar results (data not shown), demonstrating that none of these parameters (ie, type of gel, gel stain, or protein size) were detrimental to peptide yields.

Effects of Sample Handling on Adsorptive Losses of Extracts

In the experiments described previously, extracts were transferred directly to scintillation vials to minimize adsorptive losses and accurately represent the extent of peptide extraction. However, in practice, the investigator can expect much lower actual yields because of adsorptive losses associated with downstream handling. Several approaches were taken to assess downstream adsorptive losses. Initially, small aliquots of each extract were transferred to scintillation vials, and the remaining volume was measured and transferred to a microfuge tube, a fraction of the pooled extracts was counted, and the observed cpm

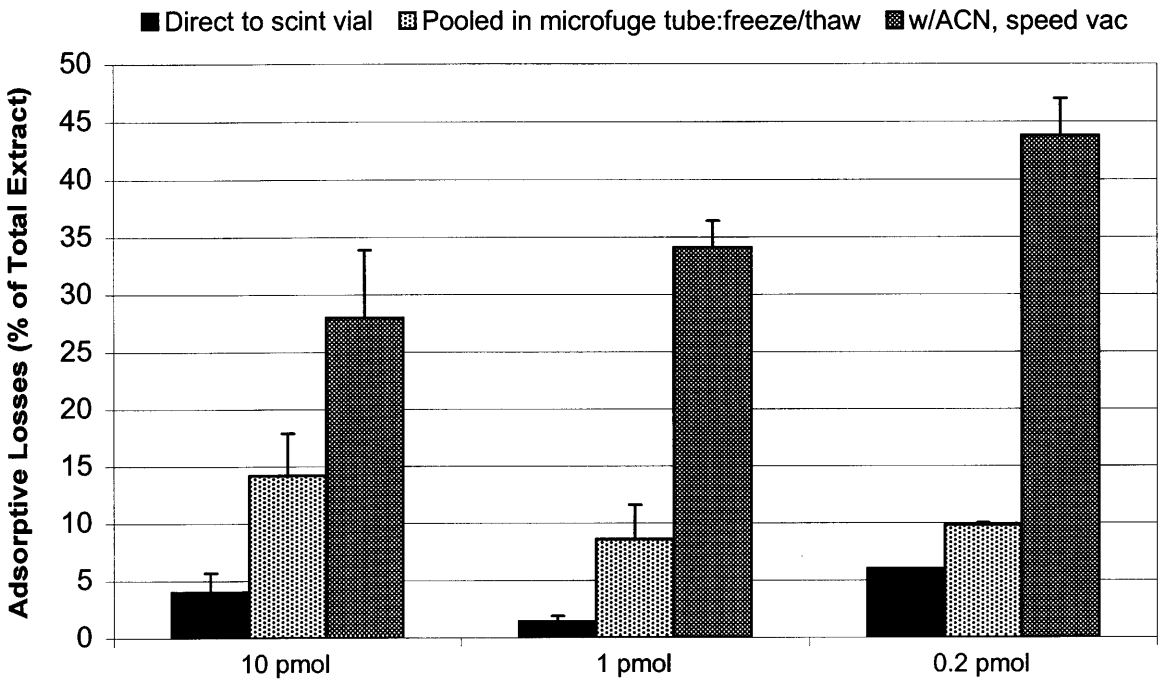

FIGURE 7

Effects of sample handling on adsorptive losses of extracts. The experiments described previously indicated the total amounts of peptides extracted but did not fully evaluate postextraction losses because extracts were transferred directly to scintillation vials. In this experiment, losses resulting from various postextraction handling methods are shown.

were corrected for the proportion analyzed. Comparisons of the amount initially extracted and the amount recovered from the total pool indicated substantial and variable losses from adsorption at 10 pmol, and greater losses occurred at lower protein amounts (data not shown). However, this approach became somewhat unreliable at 1 pmol and impractical at lower protein amounts, which represented the amounts we were most interested in evaluating, because of low radioactivity levels in small aliquots and subsequent use of large multipliers to convert to total radiolabel in the entire sample. In this regard, values determined for the second and third extractions were particularly problematic because very little total radiolabel was recovered in these fractions (Figs. 1 and 6).

In several subsequent experiments, an alternative strategy was used to more reliably measure postextraction adsorptive losses over a range of concentrations, including femtomole protein amounts. Six replicate Coomassie blue stained bands at each protein amount tested were digested and extracted in parallel using our standard method (see Methods section), except 0.1% TFA in 50% acetonitrile was used for the third extraction instead of 0.1% TFA alone for one set of two samples. Duplicate samples were then processed in several alternative downstream handling strategies (Fig. 7). When extracts were transferred directly into scintillation vials, the only site where peptides could be lost was the single pipet tip used for the transfer, which was about 5% of the total extracted amount. The sum of these two amounts (directly transferred extract and transfer pipet tip wash) was defined as 100% of the total extract for all samples at that protein amount.

Although the losses to pipet tips in this data set are not large, two points are worth noting. A single pipet tip was used for all three extracts of each sample, although typically the experimenter is more likely to use a new tip for each transfer, which would further increase losses over those estimated here. Also, the extracts were in contact with the pipet tip for only a few seconds during the transfer, indicating that peptide adsorption to plastic is a very fast reaction.

Adsorptive losses increased substantially for all protein levels when extracted tryptic peptide samples encountered more realistic sample processing procedures than a single transfer using a single pipet tip. Minimal practical processing of extracted peptides from experimental samples is represented by the second data series, in which the extracts were pooled, frozen overnight, thawed, and then counted. This condition approximates handling of a sample that

does not require concentration or removal of organic solvent (acetonitrile was limited to the 10% acetonitrile in the trypsin digestion buffer and first extract) before injection onto an HPLC or LC-MS, or application to a MALDI MS target.

Another common sample processing strategy is represented by the third data series in Figure 7. These samples were processed using a third extraction containing 50% acetonitrile. The sample was then concentrated to one third of the original volume in a Speedvac to partially concentrate the sample and to remove the acetonitrile. This sample handling method resulted in the largest losses, which increased with decreasing protein amounts. Several other experiments using Speedvac drying showed the same trends as the experiment shown in Figure 7, although the extent of peptide losses varied in a somewhat unpredictable manner (data not shown). In general, the greatest peptide losses appeared to occur when samples were nearly or completely dried as opposed to reducing the volume by 50% to 75%, although this trend was not universal. In some experiments, especially when in-gel alkylation was combined with a final extraction containing 50% acetonitrile, the total amount of peptide extracted increased slightly (about 5% to 10%), but these gains were consistently negated by the increased losses associated with Speedvac concentration. Overall, losses using Speedvac drying were usually at least 25% and ranged up to 50% or more of the total extraction solution, especially for low-level samples.

DISCUSSION

Our goal in this study was to systematically evaluate peptide losses and recoveries throughout the in-gel trypsin digestion process and to use these results to develop optimized robust in-gel methods that should improve peptide yields in the currently challenging mid- to low-femtomole range. Although it is well documented that peptide yields decline as protein amounts in the gel decrease,[16,21,22] the specific steps and parameters of in-gel digestions that influence peptide recoveries at all protein levels, including low picomole amounts, remain poorly understood.[11,12]

Current in-gel digestion protocols seem to be nearly as numerous as the number of laboratories reporting their use, although most methods still resemble the original in-gel digestion methods.[7,8,10] However, the relative efficacy of alternative modifications is usually difficult to assess from published reports. Indeed, it seems that nearly every permutation of individual steps has been used by someone with some success, usually at least at high- or mid-femtomole levels of protein in the gel and sometimes at the 100- to 200-fmol level. These permutations include different gel types, stains, and sizes; mincing or grinding up the gel slice or leaving the gel piece intact; high or low trypsin concentrations; alkylation or no alkylation; different digestion buffers; small trypsin volumes completely adsorbed by the gel or larger volumes that immerse the gel; short or overnight digestion times; variable numbers and volumes of extractions; optional use of high acetonitrile concentrations to extract peptides; and vacuum concentration or drying of samples. Two related reasons for the proliferation of in-gel digestion methods appear to be: (1) the relatively nonquantitative nature of mass analysis methods and (2) their sensitivity to contaminants that make direct evaluation of improvements in peptide yields difficult and the excellent sensitivity of current instruments that can produce positive results even when poor or variable peptide yields occur. Despite this latter advantage, further improvements in peptide yields at limiting levels should lead to improved practical detection sensitivity under otherwise identical conditions.

The use of two metabolically radiolabeled unrelated proteins prepared from a single fusion protein allowed us to systematically evaluate peptide yields without the potential complication that labeled and unlabeled peptides might behave differently, as can occur with harsh iodination methods. At the 1-to 10-pmol level, peptide yields that could be extracted and transferred directly to a scintillation vial were extremely good, typically 80% to 90% of the amount in the original gel band. This high yield was essentially unaffected by alkylation or nonalkylation, gel type (ie, Tricine, NuPAGE, or Tris-Glycine gels), or stain (ie, Coomassie blue, Ruby Red, or no stain). In addition, preliminary experiments at the 100- to 200-fmol level show that the extractable peptide yields remain high and are usually more than 70%. These data demonstrate that further improvements in peptide extraction are probably not feasible and in any event would not be very beneficial. The high amount of peptides recovered in the first extraction indicates that multiple extractions or high volume extractions (>25 to 50 μL) are of minimal benefit if gel volumes are small (<5 to 10 mm^3). More importantly, it is clear that the biggest problems are postextraction processing (Fig. 7) and possibly microheterogeneity of some resulting peptides (Fig. 3).

The potential for peptide adsorption to surfaces is well recognized, and several strategies have been advanced to minimize losses, including use of low levels of detergents[22,23] or adsorption of peptides onto solid matrices.[16,24] However, two steps that are common in many versions of current in-gel digestion

methods are the use of acetonitrile for peptide extraction and Speedvac concentration of samples to remove the acetonitrile or to concentrate the protein. Our results suggest that the small additional benefit of improved peptide extraction when acetonitrile extractions are used is outweighed by increased adsorptive losses during vacuum drying. An additional concern is that recoveries using the same protein under similar conditions varied appreciably within an experiment and more extensively between experiments. This variability probably largely results from subtle variations in how peptides dry as thin films on the sides of the microfuge tubes and the efficiency of redissolving this invisible film—ie, the composition of the solvent used and how effectively this small volume is passed over the region of the tube that contains the dried film of peptides. This interpretation is supported by the observation that we could slightly improve recoveries by extensively and carefully pipetting the remaining third of the sample volume over the entire region of the tube that contained the original volume without spreading the peptides to new regions of the microfuge tube. Another important consideration is that losses at this step increased markedly at lower peptide concentrations, suggesting that vacuum drying should be avoided when femtomole levels of proteins are being analyzed. The deletion of the acetonitrile and vacuum drying has the added benefit of simplifying the manual procedure, and their deletion is particularly advantageous for automating the in-gel digestion process.

Another potentially important consideration is the type of gel, the gel thickness, and the gel volume when gel bands are not minced. Presumably gel thickness and gel band dimensions would not be a factor if bands are minced into small pieces (eg, 1 mm^2), as recommended in some methods. However, use of unminced bands simplifies processing of large numbers of samples and minimizes the contamination of extracts with minute gel particles. Surprisingly, peptide yields were low when 0.5-mm gels were used, and this reduced yield appeared to be largely caused by loss of protein from the gel during the Coomassie blue fixing and staining step and possibly destaining of the intact gels. In retrospect, increased losses from thin gels during staining should not be surprising because all proteins are initially saturated with SDS and molecules near the gel surface can freely diffuse out of the gel during the first few minutes of the staining step before the SDS is dissociated from the proteins. Because different proteins are likely to release SDS at different rates and small proteins usually diffuse more rapidly, the use of thin gels may lead to more variability in sample recovery at this predigestion stage. Overall, 1.0-mm-thick gels appear to be ideal if gels are not minced because recoveries are high, and regardless of the size of the gel pieces, the maximum distance that trypsin must diffuse in and peptides must diffuse out is 0.5 mm (one half of the gel thickness).

Two related factors to gel thickness are gel band volume and protein density in the gel. The use of larger gel volumes and 1.5-mm-thick gels in this study did result in reduced peptide yields compared with compact bands from 1.0-mm gels, although the basis for this reduced yield is not clear (Fig. 5). Nonetheless, 1.5-mm gels and larger volumes are not optimal because diffusion out of the gel is less effective during the first extraction, necessitating multiple and larger volume extractions. In addition, preliminary experiments at lower protein concentrations than those used here show impaired proteolytic efficiency (data not shown). Reduced enzyme activity at low concentrations has been previously described as a major limitation of low level in-gel digestion.[16]

In summary, there appear to be two major factors that limit peptide recoveries at femtomole levels, adsorptive losses of initially extracted peptides and reduced protease activity at low substrate concentrations.[16] The results of this study suggest an optimized in-gel trypsin digestion strategy in which proteins are separated on 1.0-mm-thick SDS gels. When 1D gels are run, the use of mini-gels with narrow wells usually maximizes protein density in stained bands. In addition, gradient gels are sometimes useful for sharpening bands, especially if multiple proteins of interest covering a wide range of molecular weights are to be separated. Similarly, the second dimension of 2D gels should optimally be 1.0 mm thick, although the use of mini-gels is usually not critical because low abundance spots typically have very small diameters even on full-sized 2D gels. Either 1D or 2D gels can then be stained with Coomassie blue or the higher-sensitivity Ruby Red stain. Excised gel bands are digested overnight with modified trypsin and then extracted one or two times with small volumes of aqueous buffer. It is most critical to minimize subsequent handling steps and exposure to plastic surfaces, and concentration by vacuum drying should be avoided. Further studies are pursuing the effects of different extraction volumes and strategies and alternative high-sensitivity stains on peptide yields and peptide mass determination.

ACKNOWLEDGMENTS

We would like to thank Peter Hembach for performing the gel analyses and determining radioactivity in gel slices. This work was supported by NIH grants CA77048, CA74294, CA10815, and HL38794.

REFERENCES

1. Wilkins MR, Sanchez JC, Gooley AA, Appel RD, Humphery-Smith I, Hochstrasser DF, et al. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev* 1996;13:19–50.
2. Wilkins MR, Sanchez JC, Williams KL, Hochstrasser DF. Current challenges and future applications for protein maps and post- translational vector maps in proteome projects. *Electrophoresis* 1996;17:830–838.
3. Anderson NL, Anderson NG. Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis* 1998;19:1853–1861.
4. Jungblut PR, Zimny-Arndt U, Zeindl-Eberhart E, Stulik J, Koupilova K, Pleissner KP, et al. Proteomics in human disease: cancer, heart and infectious diseases. *Electrophoresis* 1999;20:2100–2110.
5. Aebersold RH, Leavitt J, Saavedra RA, Hood LE, Kent SB. Internal amino acid sequence analysis of proteins separated by one- or two-dimensional gel electrophoresis after in situ protease digestion on nitrocellulose. *Proc Natl Acad Sci USA* 1987;84:6970–6974.
6. Matsudaira P. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J Biol Chem* 1987;262:10035–10038.
7. Kawasaki H, Emori Y, Suzuki K. Production and separation of peptides from proteins stained with Coomassie brilliant blue R-250 after separation by sodium dodecyl sulfate–polyacrylamide gel electrophoresis. *Anal Biochem* 1990;191:332–336.
8. Rosenfeld J, Capdevielle J, Guillemot JC, Ferrara P. In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Anal Biochem* 1992;203:173–179.
9. Fernandez J, DeMott M, Atherton D, Mische SM. Internal protein sequence analysis: enzymatic digestion for less than 10 micrograms of protein bound to polyvinylidene difluoride or nitrocellulose membranes. *Anal Biochem* 1992;201:255–264.
10. Hellman U, Wernstedt C, Gonez J, Heldin CH. Improvement of an "in-gel" digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing. *Anal Biochem* 1995;224:451–455.
11. Williams K, Hellman U, Kobayashi R, Lane W, Mische S, Speicher D. Internal protein sequencing of SDS PAGE–separated proteins: a collaborative ABRF Study. In Marshak D (ed): *Techniques in Protein Chemistry, VIII*. San Diego: Academic Press, 1997:99–109.
12. Mische S, Speicher D, Hellman U, Williams K. Internal protein sequencing. In *Encyclopedia of Bioprocess Technology*. New York: John Wiley & Sons, 1999:2100–2103.
13. Shevchenko A, Jensen ON, Podtelejnikov AV, Sagliocco F, Wilm M, Vorm O, et al. Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc Natl Acad Sci USA* 1996;93:14440–14445.
14. Jonscher KR, Yates JR 3rd. The quadrupole ion trap mass spectrometer–a small solution to a big challenge. *Anal Biochem* 1997;244:1–15.
15. Ekstrom S, Onnerfjord P, Nilsson J, Bengtsson M, Laurell T, Marko-Varga G. Integrated microanalytical technology enabling rapid and automated protein identification. *Anal Chem* 2000;72:286–293.
16. Quadroni M, James P. Proteomics and automation. *Electrophoresis* 1999;20:664–677.
17. Ursitti JA, Kotula L, DeSilva TM, Curtis PJ, Speicher DW. Mapping the human erythrocyte beta-spectrin dimer initiation site using recombinant peptides and correlation of its phasing with the alpha-actinin dimer site. *J Biol Chem* 1996;271:6636–6644.
18. Schagger H, von Jagow G. Tricine–sodium dodecyl sulfate–polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* 1987;166:368–379.
19. Sechi S, Chait BT. Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification. *Anal Chem* 1998;70:5150–5158.
20. Steinberg TH, Chernokalskaya E, Berggren K, Lopez MF, Diwu Z, Haugland RP, et al. Ultrasensitive fluorescence protein detection in isoelectric focusing gels using a ruthenium metal chelate stain. *Electrophoresis* 2000;21:486–496.
21. Courchesne PL, Luethy R, Patterson SD. Comparison of in-gel and on-membrane digestion methods at low to sub-pmol level for subsequent peptide and fragmention mass analysis using matrix-assisted laser-desorption/ionization mass spectrometry. *Electrophoresis* 1997; 18:369–381.
22. Arnott D, O'Connell KL, King KL, Stults JT. An integrated approach to proteome analysis: identification of proteins associated with cardiac hypertrophy. *Anal Biochem* 1998;258:1–18.
23. Erdjument-Bromage H, Lui M, Lacomis L, Grewal A, Annan RS, McNulty DE, et al. Examination of micro-tip reversed-phase liquid chromatographic extraction of peptide pools for mass spectrometric analysis. *J Chromatogr A* 1998;826:167–181.
24. Gevaert K, Demol H, Sklyarova T, Vandekerckhove J, Houthaeve T. A peptide concentration and purification method for protein characterization in the subpicomole range using matrix assisted laser desorption/ionization–postsource decay (MALDI-PSD) sequencing. *Electrophoresis* 1998;19:909–917.